УДК 004.4

М.В. Гиль, Ю.И. Филатов, А.М. Фонотов
ГВУЗ «Донецкий национальный технический университет», г. Донецк, Украина
Украина, 83050, пр. Б. Хмельницкого, 84, mary_kukla@mail.ru

Разработка системы предварительного отбора кандидатов на собеседование на основе методов интеллектуального анализа данных

M.B. Hil, Y.I. Filatov, A.M. Fonotov
SHEI «Donetsk National Technical University», Donetsk, Ukraine
Ukraine, 83050, etc. Khmelnytsky, 84, mary_kukla@mail.ru

Develop a System of Pre-selection of Candidates for Interview Based on Data Mining Techniques

М.В. Гіль, Ю.І. Філатов, А.М. Фонотов
ДВНЗ «Донецький національний технічний університет», м. Донецьк, Україна
Україна, 83050, пр. Б. Хмельницького, 84, mary_kukla@mail.ru

Розробка системи попереднього відбору кандидатів на співбесіду на основі методів інтелектуального аналізу даних

В статье рассматривается проблема: «Отбор ограниченного количества претендентов на вакантную должность для собеседования с работодателем. Разработка или модификация интеллектуальных методов анализа слабо сформированных данных», – и описаны возможные пути ее решения с помощью методов Data Mining.
Ключевые слова: претендент, кандидат, персонал, должности, работодатели, вакансии, должностная инструкция, Data Mining.

The article deals with the problem: «The selection of limited number of applicants for a vacant position for an interview with the employer. Development or modification of intelligent methods for weakly generated data» and describes possible solutions with Data Mining's methods.
Key Words: applicant, candidate, personnel, office, employers, jobs, job description, Data Mining.

У статті розглянута проблема: «Відбір обмеженої кількості претендентів на вакантну посаду для співбесіди з роботодавцем. Розробка чи модифікація інтелектуальних методів аналізу слабо сформованих даних», і описані можливі шляхи її вирішення за допомогою методів Data Mining.
Ключові слова: претендент, кандидат, персонал, посади, роботодавці, вакансії, посадова інструкція, Data Mining.

Введение

Постановка проблемы исследования. При приеме сотрудников работодателю необходимо принимать важное решение, от которого в перспективе будет зависеть рабочая атмосфера коллектива и эффективность функционирования всего предприятия. Поэтому в процессе решения этой задачи, наряду с определением профессионального уровня кандидата, особое внимание уделяется личностным качествам претендента: пунктуальность, ответственность, умение работать в команде, дружелюбность, ис-

полнительность и т.д. В случае неправильного выбора специалистов работодатели рискуют понести неоправданные материальные затраты и потери рабочего времени.

Для автоматизации процесса управления подбором персонала невозможно разработать универсальную формализованную модель в силу специфических требований к претендентам на рынке труда. В то же время повысить эффективность управления этим процессом можно, если создать систему поддержки принятия решений (СППР), задача которой будет заключаться в установлении соответствия уровня подготовки специалиста требованиям рынка труда.

Анализ литературных источников.

Анализ уже существующих систем, которые решают задачу отбора претендентов на собеседование, например, «Бизнес-аналитика» SAP Business Intelligence [1], показал, что наиболее эффективными решениями из используемых являются решения с применением методов интеллектуальной обработки данных – Data Mining.

Термин Data Mining часто переводится как извлечение информации, средства поиска закономерностей, анализ шаблонов. Понятие «обнаружение знаний в базах данных» (Knowledge Discovery in Databases) можно считать синонимом Data Mining [2]. Интеллектуальный анализ данных позволяет анализировать данные различных объёмов с разных точек зрения, классифицировать их, и установить зависимости между ними [3]. Применение этих методов дает преимущество в тех областях, где накоплены значительные объемы информации. Одной из таких областей является подбор персонала.

Постановка задачи

Целью данной статьи является анализ возможных подходов построения системы отбора ограниченного числа претендентов на собеседование с работодателем на основе анализа анкет претендентов.

Задачи научные.

1 Определение критериев важности профессиональных данных кандидата на основе интеллектуального анализа анкет соискателей.

2 Разработка требований для создания интеллектуальной системы отбора кандидатов.

На основе анализа предметной области можно выделить следующие задачи, которые должна решать система подбора кандидатов:

1 подготовка данных из анкет кандидатов к виду, пригодному для их дальнейшего анализа;

2 предварительный отбор анкет по выдвинутым требованиям и ограничивающим факторам;

3 анализ отобранных анкетных данных;

4 дополнительное тестирование кандидатов по адаптивным сценариям;

5 отбор кандидатов на собеседование.

Для выделения кандидатов на собеседование необходимо первоначально определить критерии, по которым можно производить отбор, и степень важности каждого критерия.

Рассмотрим более подробно процесс отбора кандидатов. Предположим, имеется список претендентов на вакантную должность, который характеризуется набором данных: ФИО, дата рождения, пол, информация о профильном образовании и стаже работы по специальности, профессиональными характеристиками, результатами индивидуального психологического тестирования и т.д. Кроме этого есть техническое задание на подбор персонала определенной специализации и уровня подготовки.

Формально эти данные можно представить в следующем виде:

1 Задано множество $\mathbf{S^m}$ описаний требований, предъявляемых к претенденту:

$$\mathbf{S^m=\{S_i^m\};\ S_{ij}^m=\{Q^m\};\ S^m=\{Q_t^m\}; i=1, I^m;\ j = 1, J^m;\ t=1, Q^m_{max};\ m=1,M;}$$

$\mathbf{I^m}$ – число описаний требований к претендентам на вакансии **m**-о типа;

$\mathbf{J^m}$ – число признаков описаний требований к претендентам на вакансии **m**-о типа;

$\mathbf{Q^m_{max}}$ – число различных значений признаков описаний требований к претендентам на вакансии **m**-о типа.

2 Множество $\mathbf{G^m}$ коэффициентов – «критериев важности», определяющих «уровень важности» всех значений признаков.

$$\mathbf{G^m = \{G_t^m\}.}$$

3 Множество $\mathbf{L^m}$ описаний программных средств-претендентов каждого из **m** типов:

$$\mathbf{L^m = \{L_n^m\};\ L_{nj}^m = \{Q^m\};\ Q^m=\{Q_t^m\};\ n = 1, N^m.}$$

Требуется:

Выбрать из множества $\mathbf{L^m}$ описаний претендентов на вакансии **m**-о типа требуемое количество описаний претендентов, в соответствии с предъявляемыми к ним требованиям [4].

Совокупность **P** выбранных претендентов на вакансии **m**-о типа:

$$\mathbf{P = \{P^m\};\ m = 1, M.}$$

Одной из основных задач для построения системы отбора кандидатов, является задача определения **м**ножества $\mathbf{G^m}$ – «критериев важности».

Рассмотрим возможность использования уже накопленных за время работы агентства подбора персонала знаний для выделения ранжирования критериев.

Используя алгоритм, приведенный на рис. 1, можно классифицировать все имеющиеся анкеты.

Входная информация:

– список соискателей;

– сведения из анкет кандидатов.

Выходная информация:

– список кандидатов, претендующих на вакантную должность, с указанием всех необходимых характеристик и результатов предварительного тестирования;

– отчёты с дополнительной информацией по кандидатам, сформированные и отсортированные по заданным критериям.

Анализируя выделенные классы, определим перечень атрибутов кандидатов (согласно их анкетным данным и данным тестирования), которые оказывают наибольшее влияние на отбор кандидата для собеседования, на отказ кандидату, на принятие кандидата на работу.

– Ф.И.О. (не анализируется);

– данные паспорта, ИНН (не анализируется);

– вид деятельности (анализируется на этапе выборки кандидатов из базы данных агентства);

– адрес;

– дата рождения;

– общий трудовой стаж работы;

– стаж работы на вакантной должности;

– пол;

– специальность и квалификация;

– занимаемые должности и должностные обязанности с трёх последних мест работы;

– причины увольнения;
– наличие стажа работы на руководящих должностях;
– опыт работы с компьютером;
– прохождение собеседования;
– успешность выполнения должностных обязанностей;
– языковые навыки и т.д.

Каждый атрибут в анкете соискателя имеет свой «критерий важности» – число от 1 до 20.

Для обработки имеющихся данных методами Data Mining необходимо произвести кодирование данных из анкет кандидатов на собеседование. Приведем пример такого кодирования для наиболее важных атрибутов соискателя.

Возраст (рассчитывается на основании даты рождения)

От 18 до 23 лет: 0.
От 24 до 29 лет: 1.
От 30 до 39 лет: 2.
От 40 до 54 лет: 3.
От 55 до 65 лет: 4.
Больше 65 лет: 5.

Общий трудовой стаж работы (в годах)

[0; 0,5): 0.
[0,5; 1): 1.
[1; 3): 2.
[3; 5): 3.
[5; 10): 4.
≥10: 5.

Стаж работы на вакантной должности (в годах)

[0; 0,5): 0.
[0,5; 1): 1.
[1; 3): 2.
[3; 5): 3.
[5; 10): 4.
≥10: 5.

Пол

Женский: 0.
Мужской: 1.

Специальность

Соответствует требуемой: 0.
Не соответствует требуемой: 1.
Родственная специальность: 0,6 (например, степень родства).

Занимаемые должности

Хотя бы одна соответствует требуемой: 0.
Ни одна не соответствует требуемой: 1.

Причины увольнения

По инициативе работника: 0.
По инициативе собственника (администрации) предприятия: 1.
По обоюдному согласию работника и администрации: 2.
Другое: 3.

Опыт работы с компьютером

Да: 0.
Нет: 1.

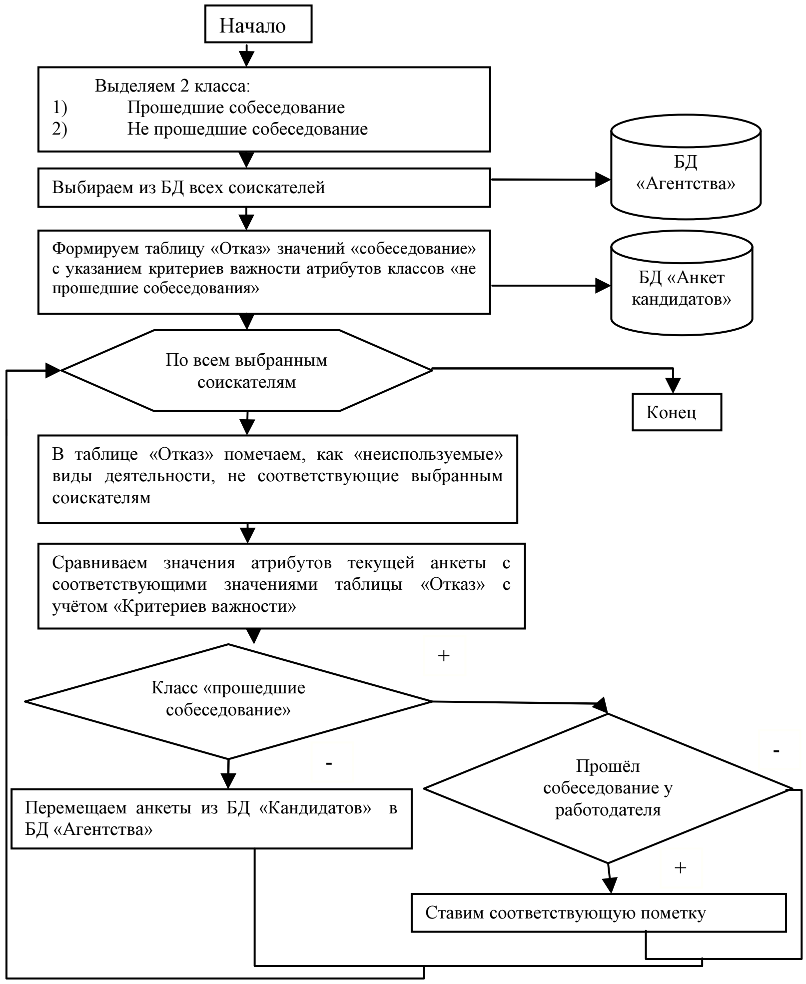


Рисунок 1 – Обобщённый алгоритм работы системы предварительной обработки БД агентства

Такой подход позволит путем анализа имеющейся базы данных выделить наиболее важные критерии отбора кандидатов на собеседовании, выделить скрытые зависимости, формализовать процесс отбора кандидатов на собеседование, сформировать базу знаний для дальнейшего развития системы отбора кандидатов на собеседование.

Выводы

В ходе написания данной статьи были сформулированы требования для разработки системы предварительного отбора кандидатов на основе методов интеллектуального анализа данных с использованием системы поддержки принятия решений, в основу которой должна быть положена интеграция методов анкетирования и тестирования.

Рассматриваемая система будет полезна для работодателей, которые хотят найти квалифицированных работников на вакантные места, а так же для работников, которые хотят получить или сменить работу.

Литература

1. SAP Business Management Software Solutions, Applications and Services [Электронный ресурс]. – Режим доступа : http://sap.com.

2. Файяд. Достижения в области обнаружения знаний и Data Mining (глава 1) AAAI / Файяд, Пиатетски-Шапиро, Смит, Узурузами. – M IT Press, 1996.
3. Андерсон Градюэйт. Data Mining. Технология, подготовленная для управления 274A / Андерсон Градюэйт. – Билл Палац, Весна, 1996.
4. Кретов В.С. Компьютерный метод отбора персонала фирмы / В.С. Кретов, И.М. Коробицын // Науковедение интернет-журнал [Электронный ресурс] Режим доступа: http://naukovedenie.ru/index.phpid=160.
5. Gil M.V. Data mining as part of the information technologies market of the human resources / M.V. Gil, A.M. Fonotov // Информационные управляющие системы и компьютерный мониторинг (ИУС и КМ) : мате-риалы III всеукраинской конференции студентов, аспирантов и молодых учёных. – Донецк : ДонНТУ, 2012

Literatura

1. SAP Business Management Software Solutions, Applications and Services [electronic resource] : http://sap.com.
2. Fayyad, Piatetsky-Shapiro, Smyth, and Uthurusamy Advances in Knowledge Discovery and Data Mining, (Chapter 1) AAAI/MIT Press 1996.
3. Data Mining. Technology Note prepared for Management 274A. Anderson Graduate. – Bill Palace, Spring, 1996.
4. Kretov VS Korobitsyn IM, Computer method for selecting personnel of the company // Science of Science online journal [electronic resource]: http://naukovedenie.ru/index.php?id=160
5. Gil M. V., Fonotov A. M, Data mining as part of the information technologies market of the human resources, Information control systems and computer monitoring (ICS and CM) / Materials III National Conference of Students and Young Scientists. – Donetsk, Donetsk National Technical University. – 2012.

RESUME

M.B. Gil, Y. Filatov, A.M. Fonotov

Develop a System of Pre-Selection of Candidates for Interview Based on Data Mining Techniques

The article stated requirements for the development of the pre-selection of candidates on the basis of data mining techniques with the use of a decision support system, based on integration should be based on the methods of questioning and testing.

The developed system will provide an objective assessment of relevant qualifications of applicants employer's requirements, and reduce the time required to collect and process the results of an interview with an expert candidate based on the use of intelligent information technology.

The system will be useful for employers who want to find qualified employees for vacant positions, and workers who want to get or change a job.

Статья поступила в редакцию 03.01.2013.